General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

APF76-MSFC-02222

APF76-MSFC-02222

Technical Report

INVESTIGATION OF SELECTED DISK SYSTEMS

(NASA-CR-150165) INVESTIGATION OF SELECTED DISK SYSTEMS (Teledyne Brown Engineering) 32 p HC A03/MF A01 CSCL 09B

N77-15669

Unclas

G3/60 12513

October 1976

FEB 1977
RECEIVED
NASA STI FACILITY
INPUT BRANCH

TELEDYNE BROWN ENGINEERING

Cummings Research Park • Huntsville, Alabama 35807

INVESTIGATION OF SELECTED DISK SYSTEMS
APF76-MSFC-02222

October, 1976

Prepared For

Data Systems Laboratory
George C. Marshall Space Flight Center
Huntsville, Alabama

Prepared Under
Contract No. NAS8-31488

By

Advanced Projects Division
Teledyne Brown Engineering
Huntsville, Alabama

ABSTRACT

This report examines the large disk systems offered by IBM, UNIVAC, Digital Equipment Corporation, and Data General. In particular, these disk systems are analyzed in terms of how will available operating systems take advantage of the respective disk controller's transfer rates, and to what degree all available data for optimizing disk usage is effectively employed. In the course of this analysis, generic functions and components of disk systems were defined and the capabilities of the surveyed disk system were investigated. The results of these activities are also included in this report as background information.

Approved:

R. R. Parker, Ph.D.
Manager
Data System Projects

Approved:

R S McCarter

R. S. McCarter
Senior Vice President

TABLE OF CONTENTS

1. INTRODUCTION

Direct access storage devices such as disks make multiprogramming and on-line data bases possible. Consequently, the performance of computer systems is strongly influenced by the capability of the equipment and software related to the use of disks. Many computer vendors now offer systems that include high performance, large capacity discs similar to the IBM 3330. This report presents the results of a study that examined the products of four companies, two mainframe manufacturers and two mini-computer manufacturers, to determine if discs were efficiently utilized in each system.

The four systems that were the subject of the study are:

- IBM 370 with VS1 and 3330 Disk System,
- UNIVAC 1108/1110 with EXEC 8 and 8440 Disk System,
- Digital PDP 11/70 with RSX-11D and RP05 Disk System,
- Data General Eclipse with RDOS and 4231 Disk System.

As expected, the two mainframe systems had more hardware features related to use of the disks, and they employed more sophisticated disk management software than the two minicomputer systems. Nevertheless, even the latter have fewer limitations than the advanced systems of just a few years ago.

Both hardware and software features determine the performance of a particular disk system. The currently available hardware components and associated control software offer an almost infinite set of possible system configurations and capabilities. Due to the scheduling/control complexity and the large number of possible configurations for disk

systems employing features to support multiple concurrent data paths (such as dual-ported disk units, multiple disk controllers per channel, multiple channels per disk controller, etc.) it is not possible to adequately analyze disk systems in general in this brief investigation. In addition, the functioning of such complex systems is particularly application and environment dependent. Therefore, the thrust of this investigations was to examine a well defined subset of disk systems. In particular, this investigation was restricted to the basic disk system supported by one disk controller and one I/O channel.

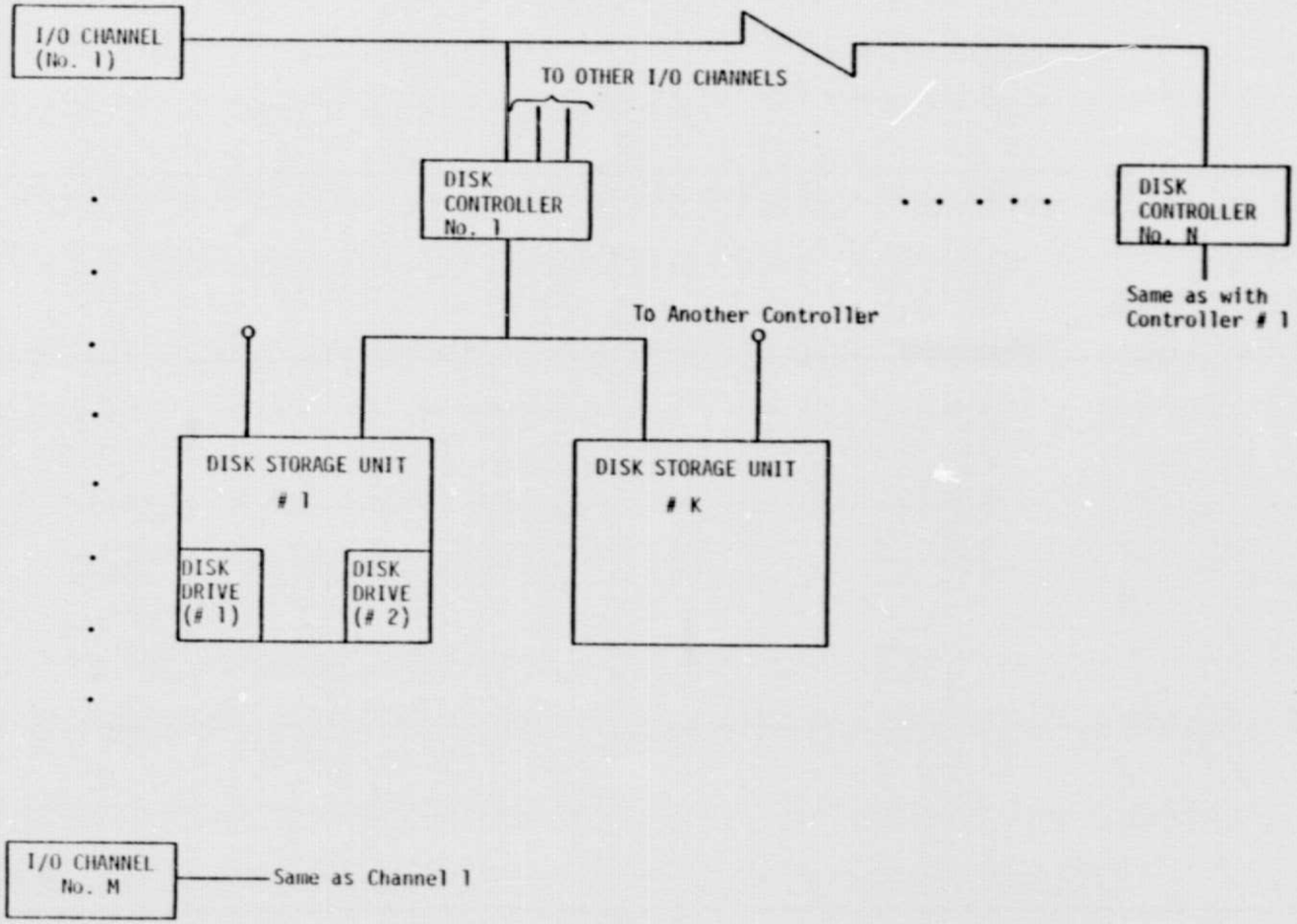
Section 2 of this report defines terminology and discusses the operation of disk systems in general. Section discusses the equipment features that influence the effective use of disks and summarizes the specifications of the disk system considered. The techniques used by each vendor's operating system to take advantage of their disk controller's capabilities are addressed in Section 4. This Section is intended to identify instances where the operating systems fail to use available system features to aid in the efficient scheduling and processing of disk data transfers.

2. DISK SYSTEM PRINCIPLES

A disk system includes several hardware modules that perform the transfer of data between a computer and the disk storage media. A representative hardware configuration is illustrated in Figure 2-1, and includes the following:

- I/O Channel to the computer: Interface electronics (possibly including a dedicated processor) and cabling for effecting direct data transfer between the disk controllers and computer memory.
- Disk Controller: Hardware device that receives and interprets command information from the I/O channel, controls the reading and writing of data on multiple disks, and generally performs error detection and correction.
- Disk Storage Unit: An equipment unit containing one or more disk drives, the read/write heads, and the associated electronics for controlling and monitoring read/write operations.
- Disk Drive: Motor and mechanisms for controlling the rotation of the disk.
- Disk Pack: Storage medium for recording data in a disk system.

Figure 2-1 illustrates the variety of potential hardware configurations possible within a disk system. For example, each disk unit contains one or two drives, and currently available hardware allows the connection of multiple controllers (generally 2) to a single disk unit. Each controller in turn can be connected to multiple I/O channels, and a single I/O channel on some computer systems can have multiple disk controllers. Thus, a single disk drive can communicate with a computer over multiple



REPRESENTATIVE DISK HARDWARE

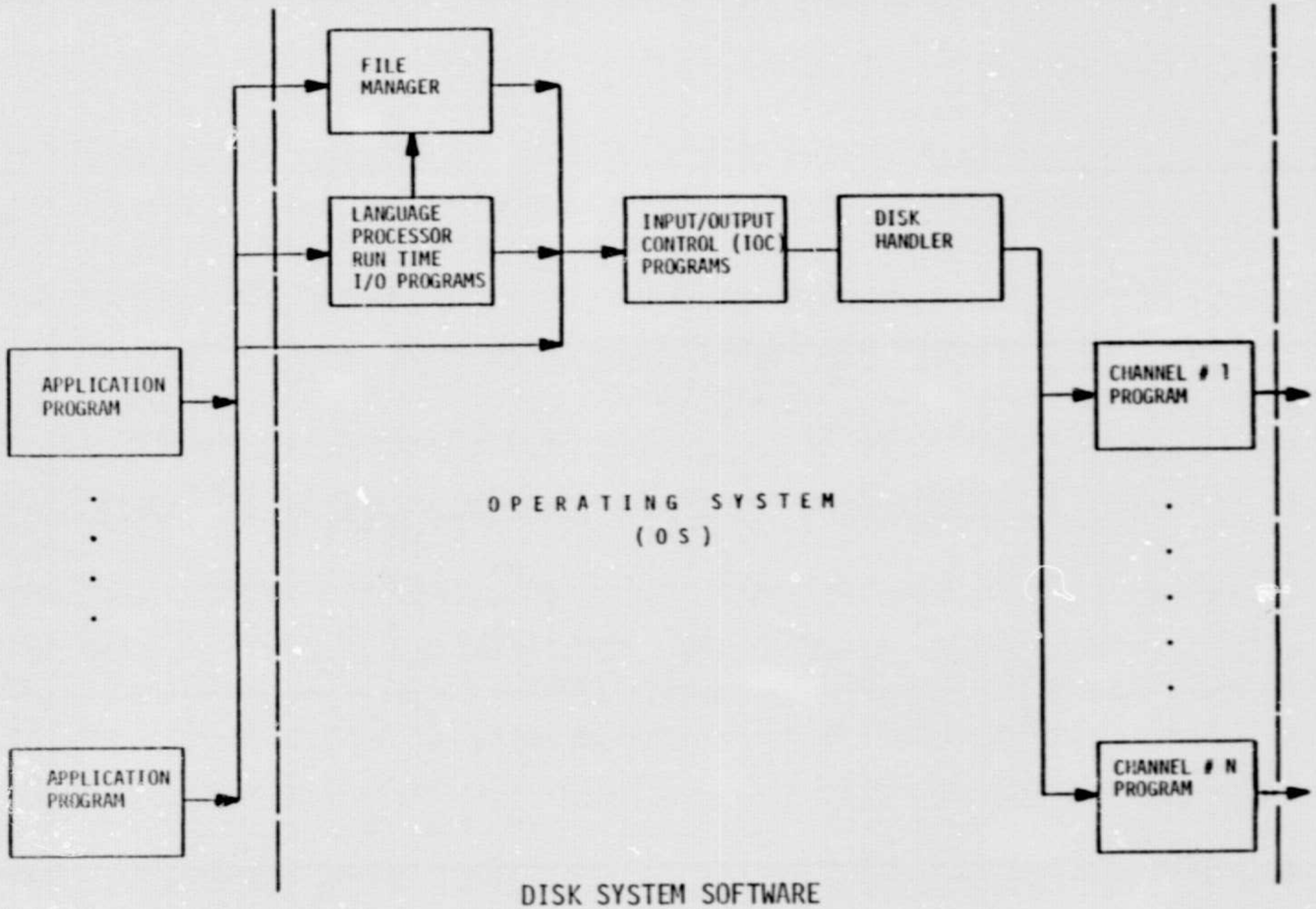
FIGURE 2-1

paths using multiple channels and/or controllers. The permissibility of concurrent activity on multiple data paths is a function of the particular paths and the capability of the computer software to schedule such operations. Many of these provisions for multiple data paths to individual disk units were developed as facilities for hardware back-up rather than for optimization of disk usage.

The software components related to the operation of a disk system are less tangible than the hardware. The basic functions can be identified but each vendor partitions these functions into modules somewhat differently. Figure 2-2 presents a generic configuration. The functions of three key components are as follows:

- Input/Output Control (IOC) Programs: Processes requests for particular physical records, orders them according to channel and priority, and dispatches them to the appropriate disk handler.
- Disk Handler: A device dependent software module that assembles or executes the I/O instructions for transmitting control commands, receiving disk status, and transferring data between I/O buffers in the computer and the disk media.
- Channel Program: One or more I/O instructions that are decoded and executed by hardware that is part of the I/O channel. Modifier data in these commands may specify device-dependent instructions for execution by the device.

The application programs, the Run-Time Language Processors and the File Manager are also shown in Figure 2-2 since they represent operating system components that transform I/O requests from the logical specification issued by an application program into the physical specification needed by the IOC program.



DISK SYSTEM SOFTWARE
FIGURE 2-2

The software component that varies the most among systems is the Channel Program. It ranges from a fairly complex set of instructions on the IBM system, which makes extensive use of the channel, to virtually nothing in some systems using different architectures.

To illustrate the interaction among modules, consider the sequence of events necessary to service a single disk request. Upon receipt of a request requiring access to a particular logical record, the File Manager uses the file description tables to construct a series of high level commands requesting specific data from a particular physical location on the disk. The IOC program dispatches the request to the appropriate handler, placing it on a request list in order of priority. The handler executes I/O commands to honor the request and/or prepares a channel program for execution by the I/O channel. Multiple commands to a single disk controller are placed in queues, either in the handler or in the controller itself. They are processed either sequentially or in response to angular position sensing data. The disk controller decodes the signals and causes the appropriate disk unit to perform the desired activity. In some systems, the disk controller maintains memory pointers and counters necessary for transfer of data between memory and the disk, while in other systems, the channel stores these values. In either case, the necessary data transfer is performed, and upon completion the data buffer is passed to the application program.

RODUCIBILITY OF THE
AL PAGE IS POOR

3. DISK SYSTEM CHARACTERISTICS AND SPECIFICATIONS

This study investigates the four computer systems identified in the introduction to ascertain if their operating systems take full advantage of the disk controller transfer rates. This section describes the features that are available for facilitating the thruput. To keep the scope of the analysis within the bounds of available effort, the investigation addresses only the performance of disk systems that consist of one I/O channel, a single disk controller, and the maximum complement of disk units for this controller.

The four disk systems examined vary widely in architecture in terms of: how the communication functions are handled; how the individual control functions are distributed between the disk storage unit, the controller, and the I/O channel; how many controllers can be connected to one I/O channel for a single access system, or to two I/O channels for a dual access system; and how many options are available for effecting control. As would be expected, the mainframe vendors offer the greatest number of options and have the most widely distributed architecture.

The primary consideration of interest in a disk system are storage capacity, transfer rate between the disks and the controller, system efficiency in terms of minimizing time when data are not being transferred between the disk and the controller, and efficiency in the management of the I/O channel. Some of these considerations are functions of hardware, and some are functions of software. The following subsections discuss the hardware features that contribute to the efficiency of each system examined.

3.1 IBM 3330 DISK SYSTEM

The IBM 3330 Disk System is available in a number of configurations with a variety of options for achieving control and capability. The configuration selected for analysis in this study is presented in Figure 3-1, and the specifications for the system are presented in Table 3-1. It incorporates the Model 3830-2 Storage Control (Other possibilities were the Integrated Storage Control and the Integrated File Adapter.) and the Block Multiplexer Channel (Another possibility was the Selector Channel.).

The disk controller functions of the 3330 disk system are allocated to the 3830-2 Storage Control and to the 3333 Disk Storage and Control, which also contains two of the eight disk drives under the control of this arrangement. Although available documentation failed to provide a detailed breakout of the functions allocated to each equipment module, characteristics of the disk controller were presented. The following features are particularly significant:

- Rotational position sensing for each drive,
- Data buffering (64 bytes) for burst transfer of data blocks on the I/O channel, and
- Use of an internal queue to stack certain instructions received from the I/O channel (up to one sequence for each disk drive).

The Model 3830-2 has a single data path that is dedicated to the selected drive from the time a Data-In or a Data-Out command is issued until the data transfer is complete. Disk seeks, angular position seeks and other

FIGURE 3-1

IBM 3330 DISK SYSTEM

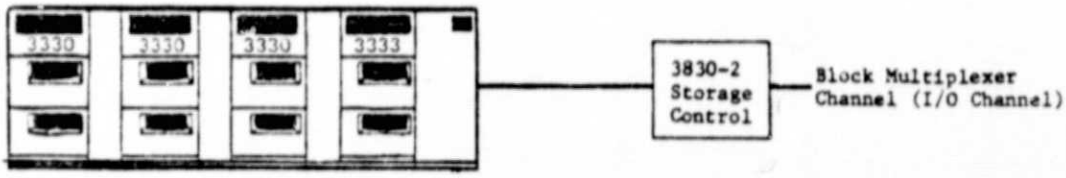


TABLE 3-1

IBM 3330, MODEL 1, DISK SYSTEM SPECIFICATIONS

CAPACITY/DRIVE	100M BYTES
No. OF DRIVES/CONTROLLER	2-8
DATA TRANSFER RATE	806K BYTES/SEC
AVERAGE HEAD SEEK TIME	30 MSEC.
AVERAGE ROTATIONAL DELAY	8.4 MSEC.
No. DISK SURFACES	19
No. TRACKS/SURFACE	411
No. SECTORS/TRACKS	NOT SPECIFIED
No. BYTES/SECTOR	NOT SPECIFIED
I/O CHANNEL SIZE	8 BITS (1 BYTE)
I/O BANDWIDTH	1.5 MEGABYTES/SEC

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

control commands, which are in the queue, can execute concurrent with a data transfer. They also are overlapped in their execution.

A Block Multiplexor Channel (BMC) was chosen for this study because its capabilities are necessary for effective use of the 3830-2 Storage Control. It includes hardware that can operate, autonomous of the CPU, to execute the channel programs that are prepared by the disk handler in response to I/O requests. Further, it can time share the execution of the channel programs. For example, after sending a seek command for disc drive # X as a result of instructions in one channel program, the I/O channel is available for requesting status of disk drive # Y, which is the next step in a different channel program. The Block Multiplexor Channel is also free in the period of time (approximately 79 μs) between the successive data blocks that are transmitted by the 3820-2 Storage Control during a Data-In or Data-Out transfer.

The 3330 disks are organized by records that are written and read independently of sector boundaries. Records are written sequentially beginning at the index of a track. Individual records may be variable in length, and a record can overflow a track by continuing to the next track on the cylinder. A sector register is associated with each disk drive, but it is used only for indicating the angular position of the disk pack relative to its index. Interrecord gaps are placed on the disk track to mark the beginning and end of records and of certain fields within a record. No sector gaps appear, however. The disk controller uses a sector address as an indication of the angular interval where it is to search for the start of a specified record. The register that contains the current angular position of the drive cannot be read by the computer.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

3.2 UNIVAC 8440 DISK SYSTEM

The UNIVAC 8440 Disk System configuration that was analyzed in this study is presented in Figure 3-2. The I/O channel shown interfaces to the central computer through on Input/Output Access Unit (IOAU) that allows the transfer of data directly between the disk and main memory. Specifications for the 8440 disk system are presented in Table 3-2. The 8440 is available in a number of configurations that exercise a variety of options for improved flexibility and performance.

Although the 8440 Disk System has considerable capabilities, it should be emphasized at this point that the system is being phased out of use by UNIVAC, and has been replaced in their line of equipment with the 8430/8433 Disk System, which has added storage capability and thruput.

The disk controller functions for the 8440 system are performed by a single unit, the Model 5033 Control Unit. It has a number of features that enhance data transfer between the disk and main storage. These features include:

- Capability to perform overlapped seek commands on all eight drives of any one controller,
- Provisions for handling overflow records without re-positioning the read/write heads,
- Provisions for angular position sensing (32 angular sectors around the circumference of the disk) and priority control to optimize data transfers,
- Data buffering that allows for up to a 5 μsec channel response time, and
- Assembly and disassembly of data words (36 bit on processor end to/from 8-bit on disk storage unit end).

FIGURE 3-2
UNIVAC 8400 DISK SYSTEM

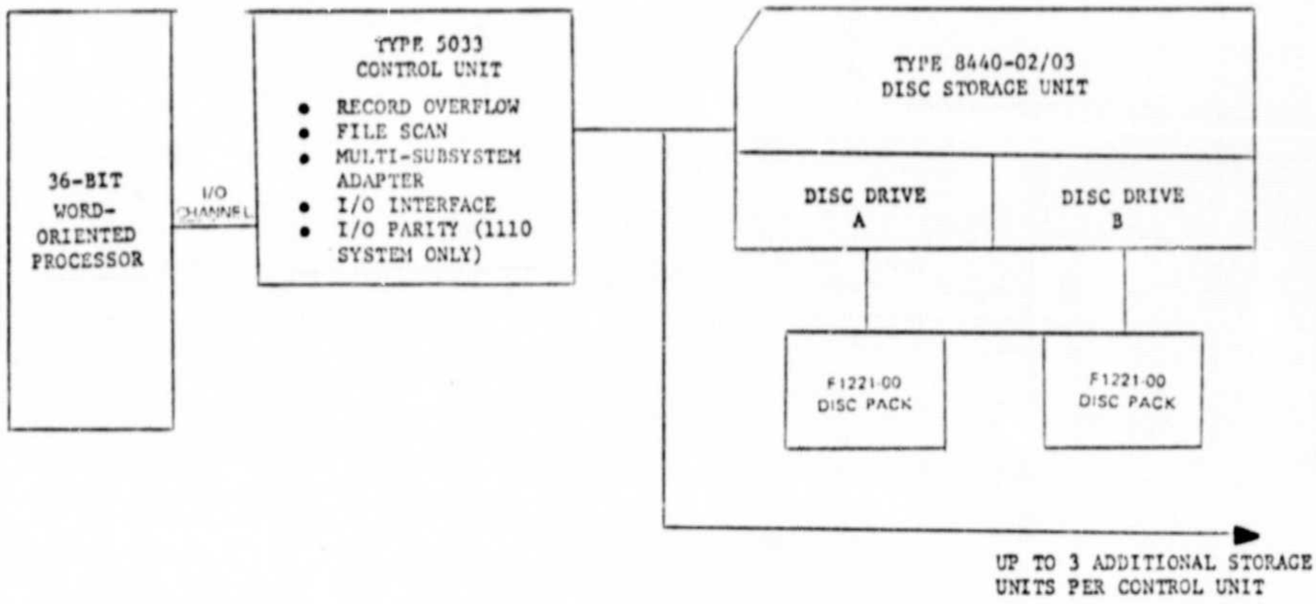


TABLE 3-2

UNIVAC 8400 DISK SYSTEM SPECIFICATIONS

CAPACITY/DRIVE	90M BYTES
No. OF DRIVES/CONTROLLER	2-8
No. OF DRIVES/CHANNEL	32
DATA TRANSFER RATE	625K BYTES/SEC
AVERAGE HEAD SEEK TIME	30 MSEC
AVERAGE ROTATIONAL DELAY	12.5 MSEC
No. OF DISK SURFACES	20
No. OF TRACKS/SURFACE	406
No. OF SECTORS/TRACK	22
No. OF BYTES/SECTOR	459
I/O CHANNEL SIZE	36 BITS
I/O CHANNEL BANDWIDTH	1.125M BYTES/SEC

The Model 5033 includes the capability to generate memory address for direct transfer of data to memory via the I/O channel. The configuration investigated for this study, however, relies upon the IOAU's capability to generate needed addresses and word counts.

The I/O channel of the 1108 is an element of the system's IOAU, which provides paths for transferring data between devices and the computer's central memory. The channel data bus is 36 bits wide and has a bandwidth of 250,000 words/sec. (Approximately 9M bits/sec. or 1.1M bytes/sec. All instructions, except those that initiate and conclude transfer of data blocks, are executed by the CPU and result in one word of either command, status or data being transmitted or received on the I/O channel. For block transfers of data, hardware in the IOAU controls the operation.

The format for storing data on the 8400 disk pack is based on a physical partition of each track into 22 equal sectors that are separated by sector gaps. A physical data record must begin at the start of a sector but can continue across sector boundaries. Records can be retrieved by sector address, by record number, or by key. Record number and key are distinct fields in the record that can be accessed by the disk controller for comparison with the same fields received with search commands via the I/O channel. The physical file format for standard EXEC 8 systems is the FASTRAND Drum Format (simulated), which has four drum sectors (28 words) packed into each disk sector (112 words). The register that contains the current angular position (sector) of the disk pack cannot be read by the computer.

3.3 DEC RP05 DISK SYSTEM

The DEC RP05 disk system may be interfaced to the PDP 11/70 computer directly via high-speed internal controllers (optional) or through an external RP04 disk controller and the UNIBUS. The internal high-speed controller configuration (Figure 3-3) was investigated in this study because it offers the most potential for optimizing the data thruput. As indicated in Figure 3-3, each RP05 storage unit contains a single drive, and up to eight drives can be connected to the same high-speed controller. Specifications for the RP05 disk system are presented in Table 3-3.

The high-speed disk controllers are an internal part of the PDP 11/70 CPU. They generate memory addresses for direct transfer of data between the disk storage units and memory. This controller, plus the control functions built into each RP05 disk storage unit, provides the following additional features:

- Capability to perform overlapped seeks and data transfer,
- First-in-first-out data buffering for smoother data channel operation (66 words),
- Ability to read data with or without header information,
- Automatic seek of cylinders in response to a read command, and
- Midtransfer seek to next cylinder following the operation on the last track and sector of the current cylinder to enhance spiral read/write operations.

When using the internal hi-speed controllers, data transfers between the RP05 System and memory occur in a Direct Memory Access Mode.

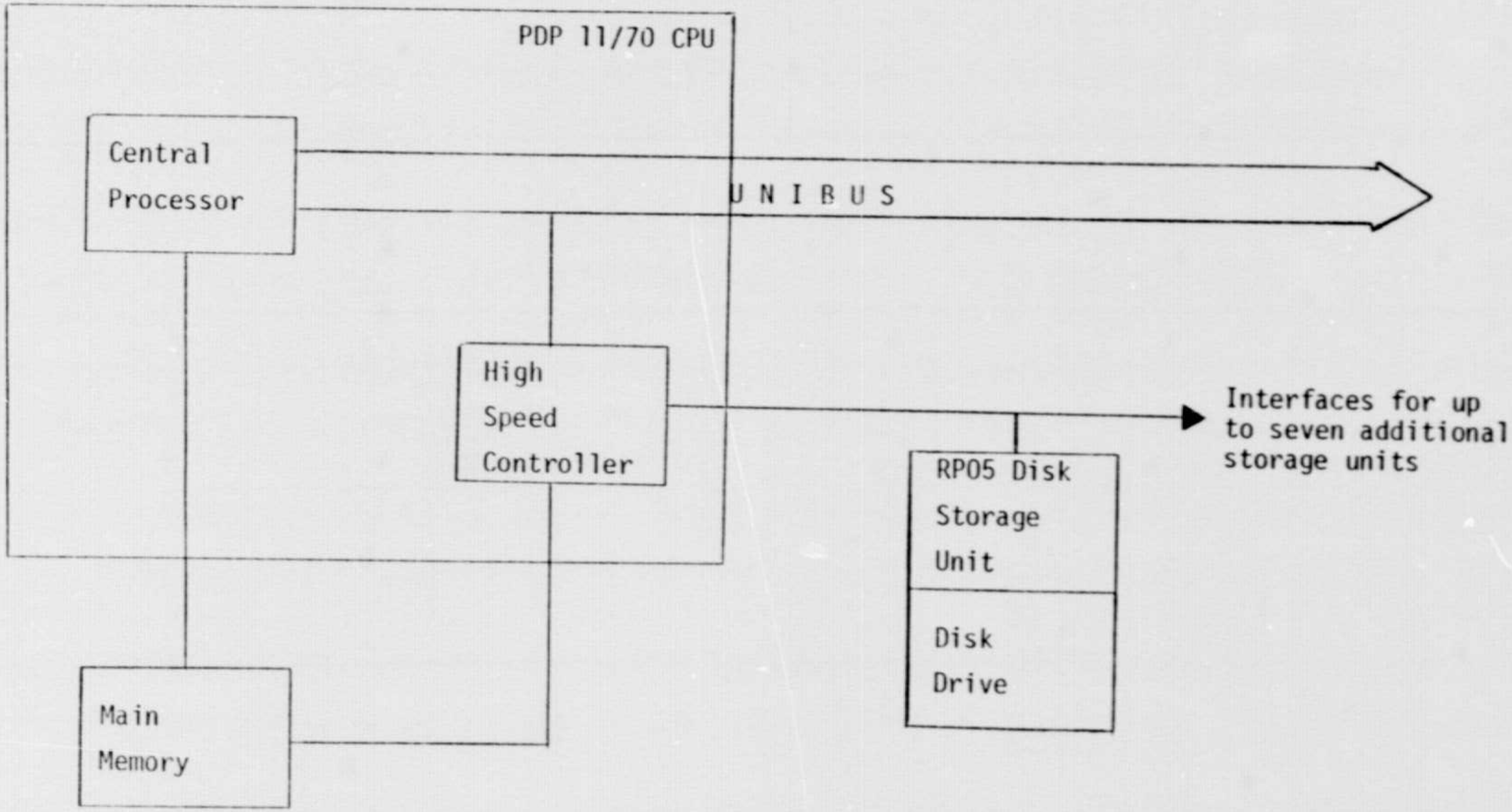


FIGURE 3-3

PDP 11/70 CONFIGURATION WITH RP05 DISKS

TABLE 3-3

DEC RP05 Disk System Specifications

Capacity/Drive	88M Bytes
No. of Drives/Controller	8
Data Transfer Rate	806K Bytes/Sec
Average Head Seek Time	28 MSEC
Average Rotational Delay	8.3 MSEC
No. Disk Surfaces	19
No. Tracks/Surface	411
No. Sectors/Track	22
No. Bytes/Sector	512
I/O Channel Size	4 Bytes
I/O Channel Speed	2.9M Bytes/Sec

An internal 32-bit wide data bus transfers four data bytes in parallel between memory and the high-speed controllers. The priority arbitration logic within the PDP-11/70 cache memory controls the timing of data transfers and updates the status of data in the cache, but the cache itself is not used for data storage. Thus, data transfers are between main memory and the RP05. The UNIBUS is used only to transmit control and status information, and as a path for interrupt requests by the controller.

The storage format used by the RP05 is based on a partition of each track into 22 sectors with sector gaps providing separation. Physical records must start at the beginning of a sector, but can continue to successive sectors for a maximum record length of 128K bytes. Each sector has a reserved field for storing the cylinder track and sector

number. The disk controller can retrieve records based on these fields, but it has no features that support record retrieval based on fields (record number, key) within the physical record. A unique feature of the RP05 recording format is that when data blocks are larger than a single sector, they are continued in every fourth sector rather than contiguous sectors. The disk system has a register in each drive for angular position sensing. It contains the number of the sector currently under the read head for the selected track. The CPU can read this register with a read instruction.

3.4 DATA GENERAL 4231A DISK SYSTEM

The Data General 4231A disk system configuration examined for this study is illustrated in Figure 3-4. Specifications for this disk system are provided in Table 3-4.

The 4231 disk system controller interfaces to a DMA channel and can support up to four disk storage units containing one drive each. Features of this controller include:

- Generation of memory addresses for DMA transfer,
- Capability to issue data transfer commands to one disk storage unit while the others (up to three) are processing seek commands,
- Use of rotational position sensing to locate the required sector,
- Automatic track switching (within a cylinder) for data blocks crossing track boundaries, and
- Data buffering capable of tolerating a data channel latency of up to 19.8 μsec.

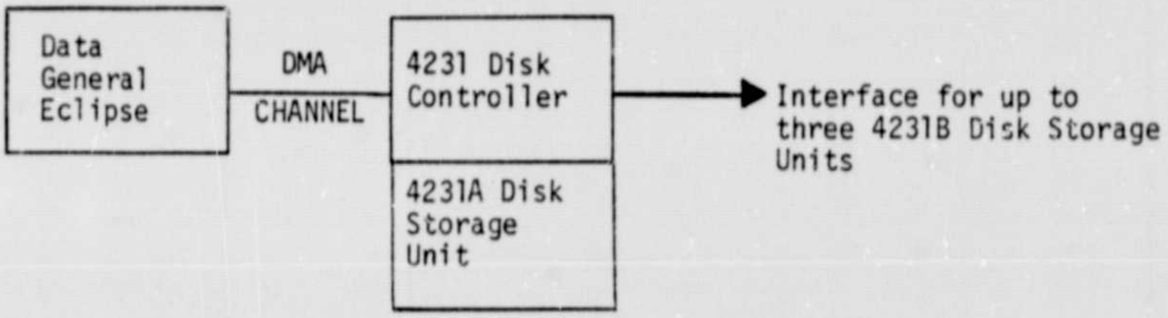


FIGURE 3-4

DATA GENERAL 4231A DISK SYSTEM

TABLE 3-4

DATA GENERAL 4231A DISK SYSTEM SPECIFICATIONS

CAPACITY/DRIVE	92M BYTES
No. OF DRIVES/CONTROLLER	4
DATA TRANSFER RATE	806K BYTES/SEC
AVERAGE HEAD SEEK TIME	30 MSEC
AVERAGE ROTATIONAL DELAY	8.4 MSEC
No. DISK SURFACES	19
No. TRACKS/SURFACE	411
No. SECTORS/TRACK	23
No. BYTES/SECTOR	512
I/O CHANNEL WORD SIZE	2 BYTES
I/O CHANNEL SPEED	1.25M BYTES/SEC (INPUT) .714M BYTES/SEC (OUTPUT)

Data transfers between the disk system and memory occur via the DMA channel. Input to the processor takes place over a 1.25 MHZ input channel. Outputs are limited to .714 MHZ by the output channel bandwidth.

The recording format on the 4231 disk system is organized with each sector having a header and a data portion. The header portion identifies each sector by cylinder, track and sector numbers. A data check field is included in each sector for error detection purposes. Data blocks larger than sectors are contained in contiguous sectors where possible. Up to 16 sectors may be written or read as a block, and all data transfers must begin on sector boundaries.

4. DISK SYSTEM PERFORMANCE

The previous section identified the characteristics and specifications of disk systems in terms of the hardware modules that comprise the system. This section is concerned with how the disk system examined manage data transfer between the disk and memory, with emphasis on the operating system interface to the disk controller. The operating systems of the computers considered in this study are too sophisticated and complex to directly assess their performance in utilizing their respective disk control units. Instead, the study examined the capabilities of the operating systems in an attempt to determine if all facilities of the disk systems were used in managing the transfer of data between the computers and the disk media. The rationale for this approach was to detect any instances where the operating systems filed to use those disk system features that could aid in effectively scheduling disk operations and data transfer.

Techniques that can be used by operating systems to optimize disk usage include:

- Issue seek commands to disk units that are not performing data transfer operations,
- Order disk requests to minimize head movement,
- Order disk requests to minimize rotational delay,
- Optimize the use of the I/O channel so as to minimize the time a disk controller waits for access to the I/O channel,

- Maintain buffers of recently read disk records so that successive requests can possibly be satisfied by data already in the buffer, thus, reducing the number of disk I/O data transfers, and
- Optimize data storage format to minimize the number of disk I/O commands needed to locate and read or write a record.

The following subsections consider the degree to which the operating systems examined use techniques such as these to optimize disk usage.

4.1 IBM 3330 UNDER VS1

Conversations with IBM technical representatives and review of system documentation led the study participants to conclude that IBM has provisions for taking advantage of all information available from the disk controller in optimizing disk utilization. Many features are standard, while others are determined by installation parameters and can be adjusted to emphasize various aspects of system performance

Under VS1, channel programs are invoked to overlap seek operations with data transfer from a drive selected for read/write of a data record. The scheme used by VS1 to order disk requests is selected by each individual system's designers, and must be specified at time of system generation. Supported ordering schemes are: first-in-first-out (FIFO), by requesting task priority, and by order of required cylinder address. Thus, designers have considerable flexibility in determining what factors of system performance will be emphasized (priority servicing versus disk utilization).

It is not possible to read angular position of disk drives on the 3330 systems; consequently, this parameter could not be a factor in

scheduling disk I/O requests. On the other hand, the disk controller can be commanded to generate an interrupt when a drive reaches a specified sector where a record starts. The feature enables other controllers to use the I/O channel until the sector is reached. VS1 does assemble channel programs to take advantage of this feature.

Because the VS1 physical file formats are strictly record oriented (e.g., not constrained by sectors), the user program and its JCL statements determine the efficiency of the data organization on the disk pack. Multiple logical records are retrieved only if they are identified to the operating system as a common physical record. Thus, there is no automatic retrieval of a certain minimum quantity of data (e.g., a sector).

4.2 UNIVAC 1108/1110 WITH EXEC 8 AND 8440 DISKS

The disk related portions of standard versions of EXEC 8 are quite basic, but a UNIVAC furnished option, called Preseek, does provide comprehensive features for efficient disk use. The following list summarizes the techniques available with EXEC 8 and Preseek to take advantage of the 8440 disk controller's transfer rate:

- Issues commands to provide concurrent seek and data transfer operations,
- Queues disk requests, within priority level, in the order requiring a minimum of total head movement time (Preseek).
- Uses sector seek commands (with interrupt enable) to free the I/O channel during the time needed for a specified sector to reach the read/write heads.

Thus, with the Preseek option EXEC 8 appears to use all available information (available from the disk controller) in scheduling disk I/O requests.

Use of the FASTRAN format may limit the performance of the disk subsystem. The study was unsuccessful, however, in deciding if the disk controller would more effectively support other formats. Because the controller supports record retrieval by record number or by key, the complete dependence on sector organization may not be necessary.

Discussion with UNIVAC technical representatives established that EXEC 8 retrieves data from all sectors that include the record being requested by an application program. Further, the most recent retrieval is examined on the next request to determine if the data fields are

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

available without a disk access. Consequently, EXEC 8 appears to make the most of sector organization by issuing commands to the disk controller only if necessary.

The use of rotational position sensing by EXEC 8 is functionally similar to that employed by VS1. The operating system activates angular position sensing by issuing a sector seek, and an interrupt signals when the angular position is located. Using this scheme, controller record searching is minimized and controller time is freed for other operations. The seek command usually is set up to select a sector ahead of the one containing the start of record. This lead time can be adjusted to the operating environment. This variable anticipation period appears to be unique to EXEC 8 among the four systems included in the study.

4.3 DIGITAL EQUIPMENT CORPORATION (DEC) PDP 11/70 WITH RSX-11D AND RP05 DISKS

RSX-11D offers a number of features for optimizating disk usage. Certain of the algorithms probably warrent further investigation and testing to establish their effectiveness. DEC representatives claim such testing was performed while developing these algorithms.

DEC's ordering of disk requests servicing, although logical, is not designed for absolute maximum througput. In particular, they honor task priorities before minimum head seek time in ordering requests. Multiple requests from a single task are, of course, ordered as issued. When DEC's optional Record Management System is in use, minimization of head seek time includes both head arm movement and rotational delay.

A unique feature of DEC's approach to disk organization is their interleaved recording format. To read or write continuous data across

sector boundaries, one track is used, three are skipped, and the sequence is repeated. For disk systems of the size considered in this study, the interleaving scheme would be a drawback since the effective transfer rate per disk would be one fourth the rate that data passes the heads. In a large system where I/O bandwidths of the computer is the limiting factor the effect of interleaving might be insignificant. DEC representatives claim that, rather than a hinderance, this recording scheme was carefully tested and selected as the most effective for the environment in which the RP05 is expected to be used, namely large file manipulations and data base management applications. Confirming this contention would require a detail of investigation that is outside the scope of this study.

Physical records are strictly tied to sectors, beginning only on sector boundaries and having no identification fields other than cylinders, track, and sector numbers. A file organization does support records longer than a sector (up to 64K words), however.

4.4 DATA GENERAL ECLIPSE WITH RDOS AND 4231 DISKS

The Data General operating system, RDOS, does not employ any sophisticated algorithms for scheduling disk I/O requests. Although I/O instructions are available for reading current head position and current angular position (sector under the read/write head); the operating system makes no use of the information available from the controller. Requests are passed to the disk handler as a first-in-first-out list. RDOS does proceed to issue seek and status commands to drives that are not transferring data; but it performs no scheduling as such.

RDOS maintains a buffer that contains the most recently read physical records. It checks this buffer prior to performing a disk read and retrieves the needed data from the buffer if it is there. If not, the needed record is read and replaces the least active record in this buffer. A user can significantly reduce the number of accesses to the disk drive by intelligently blocking several logical records into each physical record on the disk file.

5. CONCLUSION

As noted earlier, the results of this study apply to a computer system where the I/O channel bandwidth is not a limiting factor. In particular the analysis addressed systems with one disk controller and a full complement of disk drives for that controller. With these restrictions the results of the investigation support the following conclusions:

- The IBM operating system takes full advantage of the disk controller transfer rate if the user chooses the installation option that orders disk requests by cylinder address.
- The UNIVAC operating system (with the Preseek option) takes full advantage of the disk controller transfer rate except possibly for limitations on efficiency due to the use of the FASTRAN format for disk files.
- The DEC operating system does not take full advantage of the disk controllers transfer rate because although disk requests can be queued for minimum rotational delay (using the optional Record Management System), no provision is included to insure they are serviced to take advantage of disk rotational position. In addition, DEC's recording format for continuous words (every fourth sector) is likely to introduce degradation in disk utilization.
- The Data General operating system does not take full advantage of the disk controller, using neither cylinder address nor current angle position of the drive in scheduling disk requests.